Consiglio Nazionale delle Ricerche

Regularizing preconditioners based on fit techniques in the image reconstruction problem

P. Favati, G. Lotti, O. Menchi

IIT TR-18/2002

Technical report

Agosto 2002

Istituto di Informatica e Telematica

Regularizing preconditioners based on fit techniques in the image reconstruction problem.

P. Favati
IIT - CNR Via G. Moruzzi 1, 56124 Pisa, Italy.

G. Lotti
Dipartimento di Matematica, University of Parma,
Via M. D'Azeglio 85/A, 43100 Parma, Italy.

O. Menchi
Dipartimento di Informatica, University of Pisa,
Corso Italia 40, 56125 Pisa, Italy.

Abstract

Regularizing preconditioners for the approximate solution by gradient-type methods of image restoration problems with two-level band Toeplitz structure, are examined. For problems having separable and positive definite matrices, the fit preconditioner, introduced in [6], has been shown to be effective in conjunction with CG. The cost of this preconditioner is of $O(n^2)$ operations per iteration, where n^2 is the pixels number of the image, whereas the cost of the circulant preconditioners commonly used for this type of problems is of $O(n^2 \log n)$ operations per iteration. In this paper the extension of the fit preconditioner to more general cases is proposed: namely the nonseparable positive definite case and the symmetric indefinite case are treated. The major difficulty encountered in this extension concerns the factorization phase, where, unlike the separable case, a further approximation is required. Various approximate factorizations are proposed. The preconditioners thus obtained have still a cost of $O(n^2)$ operations per iteration. A large numerical experimentation compares these preconditioners with the circulant Chan preconditioner, showing often better performances at a lower cost.

1 Introduction

The image restoration problem can be discretely modelled by the linear system

$$H\boldsymbol{x} = \boldsymbol{f} - \boldsymbol{\eta}, \tag{1}$$

where $\boldsymbol{x}$ and $\boldsymbol{f}$ are n^2 vectors containing the $n \times n$ original and observed images respectively, $\boldsymbol{\eta}$ represents an unknown noise (which we will assume to be a

Gaussian white noise) and H is the blurring discrete operator. We assume that $\boldsymbol{f}$ dominates $\boldsymbol{\eta}$, otherwise the reconstruction of the original image would be impossible.

Given H and $\boldsymbol{f}$, finding a good approximation of $\boldsymbol{x}$ can be difficult, since H is generally ill-conditioned. In fact the exact solution of the system

$$H\boldsymbol{y} = \boldsymbol{f} \tag{2}$$

may differ considerably from $\boldsymbol{x}$ even if $\boldsymbol{\eta}$ is small. For this reason special techniques, known as *regularization methods*, have been devised. A widely used regularization technique suggests to solve (2) by employing the conjugate gradient method when H is positive definite or some of its generalizations for the non-positive definite case. In fact CG acts as a filtering method [19, 10]: at first the iteration reconstructs the original signal by letting only the low frequency components to pass. Successively the iteration starts to allow also increasing frequency components, corresponding to the noise. Thus the iteration must be stopped when the noise components start to interfere.

When the coefficient matrix is ill-conditioned, as in the present case, the number of iterations required by CG for obtaining a satisfactory result can be large and preconditioning is required to increase the rate of convergence. General purpose preconditioners are not satisfactory in the present case, because they are designed to reduce the number of iterations by clustering all the eigenvalues of the preconditioned matrix around 1. In this way the signal subspace, generated by the eigenvectors corresponding to the greatest eigenvalues, and the noise subspace, generated by the eigenvectors corresponding to the lowest eigenvalues, are mixed up and the effect of the noise appears before the image is fully reconstructed. In the present context a good preconditioner should reduce the number of iterations required to reconstruct the information from the signal subspace, that is, it should cluster around 1 only the greatest eigenvalues, letting the others out of the cluster.

This requires to be able to estimate a parameter $\tau > 0$, called *regularization parameter*, which separates the eigenvalues of H corresponding to the signal from those corresponding to the noise. Techniques which allow such an estimate are described in the literature (see for example [12]). They are based on the assumption that the Fourier coefficients of $\boldsymbol{\eta}$ have approximately the same magnitude for all the frequencies and they dominate the Fourier coefficients of $\boldsymbol{f}$ corresponding to the noise subspace. We assume here that an estimate of τ is available.

When matrix H, as it frequently happens, has a block Toeplitz structure, the product $H\boldsymbol{z}$ (required in the application of CG) can be computed by means of the FFT in $O(n^2 \log n)$ operations. Then the construction of the preconditioner and its use should have costs not exceeding $O(n^2 \log n)$ operations. The preconditioners based on circulant matrices satisfy this cost requirement, improve the convergence speed and can be easily adapted to cope with the noise. Unfortunately the cost of the circulant preconditioners cannot be lowered when H has a band block Toeplitz structure. In [6] a band preconditioner, called *fit preconditioner*, constructed from the symbol function of H and having regularizing effects, is introduced for positive definite matrices H with band block structure which are also separable. It costs $O(n^2)$ operations per iteration and results to be effective, giving reconstructed images with errors comparable with those obtained by the non-preconditioned CG. The number of iterations required is of the same order than the circulant preconditioners.

In this paper we want to examine the possibility to extend the application of the fit preconditioner also to the symmetric nonseparable case, in conjuction with CG if H is positive definite and MRII otherwise (MRII is a minimum residual Krylov subspace method having regularizing properties, see [9], [11]).

When H is nonseparable, the fit technique proposed in [6], which uses trigonometric basis including only cosine terms, can be no longer suitable, and sine terms should be included. The fit constructed from the symbol function of H requires to be decomposed into the product of two triangular factors. In the two-dimensional case, if H is nonseparable, factors having a finite expansion like the ones we are looking for, may not exist. Hence an approximate factorization must be considered and we propose various techniques to achieve it. The different preconditioners we obtain are then tested with a numerical experimentation, where their performance is compared also with that of a circulant regularizing preconditioner. The results of the experiments show that the fit preconditioners are very competitive in both the reconstruction efficiency and the computational cost.

In Section 2 Toeplitz matrices and related notations are introduced. The 1D version of the fit preconditioner and the 2D fit preconditioner for the separable case are recalled from [6] in Sections 3 and 4, respectively. In Section 5 several 2D fit preconditioners for the nonseparable case are described. Following a commonly applied technique, we have firstly constructed the preconditioner for a separable approximation of H. Then, operating directly from H, various 2D fit preconditioners are obtained by combining different two-dimensional fits with different approximate factorization techniques. The eigenvalues of the preconditioned matrices so obtained are analyzed in Section 6. The computational costs per iteration for the different preconditioners are listed in section 7. Finally the results of the numerical experimentation are fully reported in Section 8.

2 Preliminaries

Matrix H is defined by the so-called *point spread function* (PSF), which describes how the imaging system affects the points of the original image. In this paper we assume that the PSF is represented by a *mask* $M = (m_{i,j})$, $-w \le i, j \le w$, $w < n$. The matrix H associated with M has a 2D band $n \times n$ Toeplitz structure with bandwidth w of the form

$$H = \begin{bmatrix} H_0 & H_1 & \dots & H_{n-1} \\ H_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & H_1 \\ H_{-n+1} & \dots & H_{-1} & H_0 \end{bmatrix}, \qquad H_i = O \quad \text{for} \quad |i| > w,$$

where

$$H_i = \begin{bmatrix} h_{i,0} & h_{i,1} & \dots & h_{i,n-1} \\ h_{i,-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_{i,1} \\ h_{i,-n+1} & \dots & h_{i,-1} & h_{i,0} \end{bmatrix}, \qquad h_{i,j} = \begin{cases} m_{i,j} & \text{for } |i|,|j| \le w, \\ 0 & \text{otherwise.} \end{cases}$$

Hence the (k,j)th element of the (α,β)th block of H results to be

$$h_{k,j}^{(\alpha,\beta)} = m_{\beta-\alpha,j-k}.$$

We assume also that H is symmetric, that is $H_i = H_{-i}^T$ for $i = 0,\dots,w$. Then $m_{i,j} = m_{-i,-j}$ for $i,j = -w,\dots,w$, that is $M = JMJ$, where J is the exchange matrix of compatible size. Such a mask will be called *1-level symmetric*. If in addition all the blocks H_i are symmetric, then $m_{i,j} = m_{i,-j}$. In this case the mask, which will be called *2-level symmetric*, verifies $M = MJ$. If M is only 1-level symmetric, the quantity $\alpha_M = \|M - MJ\|_2$, which measures the asymmetry of the blocks of H, will be called *asymmetry parameter*.

In some cases the mask M is a rank one matrix, that is $M = \boldsymbol{a}\,\boldsymbol{b}^T$. Then

$$H = A \otimes B, \tag{3}$$

where A and B are $n \times n$ Toeplitz matrices whose elements are

$$a_{ij} = \begin{cases} a_{j-i} & \text{for } |i-j| \le w, \\ 0 & \text{otherwise,} \end{cases} \qquad b_{ij} = \begin{cases} b_{j-i} & \text{for } |i-j| \le w, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

From the symmetry of H it follows that $a_{-i} = a_i$ and $b_{-i} = b_i$ for $i = 1,\dots,w$. Hence M is 2-level symmetric. Matrices H satisfying (3) are called *separable*.

Finite dimensional symmetric Toeplitz matrices are generally seen as sections of bi-infinite Toeplitz matrices, generated by a *symbol* function, that is a continuous function $h : Q \to \mathbb{R}$, where $Q = [-\pi,\pi]^2$, whose Fourier coefficients are

$$m_{k,j} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} h(\theta,\eta)\ \mathrm{e}^{-\boldsymbol{i}(k\theta + j\eta)}\ d\theta\ d\eta.$$

In our case, where H is a band matrix with bandwidth w, the symbol function is

$$\begin{aligned} h(\theta,\eta) &= \sum_{k,j=-w}^{w} m_{k,j}\ \mathrm{e}^{\boldsymbol{i}(k\theta+j\eta)} \\ &= m_{0,0} + 2\sum_{k=1}^{w} m_{k,0}\cos k\theta + 2\sum_{j=1}^{w} m_{0,j}\cos j\eta \\ &+ 2\sum_{k,j=1}^{w} \Big[(m_{k,-j} + m_{k,j})\cos k\theta \cos j\eta + (m_{k,-j} - m_{k,j})\sin k\theta \sin j\eta\Big]. \end{aligned} \tag{5}$$

If the mask M is 2-level symmetric, the symbol function reduces to

$$h(\theta,\eta)=m_{0,0}+2\sum_{k=1}^{w}m_{k,0}\cos k\theta+2\sum_{j=1}^{w}m_{0,j}\cos j\eta+4\sum_{k,j=1}^{w}m_{k,j}\cos k\theta\cos j\eta. \tag{6}$$

A generalization of the classical Grenander and Szegö theorem [8] on the spectrum of symmetric Toeplitz matrices states that for any bounded function F uniformely continuous on $\mathbb{R}$ it holds

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n^2}F\big(\lambda_i(H)\big)=\frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}F(h(\theta,\eta))\,d\theta\,d\eta,$$

where $\lambda_i(H)$ are the eigenvalues of H. Moreover, if $h_{\min}$ and $h_{\max}$ are the minimum and maximum values of h respectively, with $h_{\min}<h_{\max}$, then for any n

$$h_{\min}<\lambda_i(H)<h_{\max}\qquad\text{for any } i=1,\dots,n^2.$$

In particular, if h is *positive*, then $h_{\min}>0$ and H is positive definite.

Moreover, if two symmetric 2D Toeplitz matrices H and P are given, generated respectively by the symbol functions $h(\theta,\eta)$ and $p(\theta,\eta)$, with $p(\theta,\eta)>0$ for any θ, η, the eigenvalues of the matrix $P^{-1}H$ lie between the minimum and the maximum of the function $p^{-1}(\theta,\eta)h(\theta,\eta)$ [17].

In order to construct a good preconditioner for matrix H, an approximate knowledge of the eigenvalues of H should be available. From the previous theorem the set

$$\mathcal{S}_H=\{h(\theta_k,\eta_j),\quad k,j=0,\dots,n-1\}$$

of the sampled values of h in the nodes

$$\theta_k=-\pi+k\,\delta\quad\text{and}\quad\eta_j=-\pi+j\,\delta,\quad\text{with}\quad\delta=2\pi/n, \tag{7}$$

can be assumed as an acceptable approximation of the set of the eigenvalues of H.

If H is separable as in (3), the symbol function is given by the product

$$h(\theta,\eta)=h_A(\theta)h_B(\eta),$$

where

$$h_A(\theta)=a_0+2\sum_{k=1}^{w}a_k\cos k\theta,\qquad h_B(\eta)=b_0+2\sum_{j=1}^{w}b_j\cos j\eta. \tag{8}$$

Let

$$\mathcal{S}_A=\{h_A(\theta_k),\ k=0,\dots,n-1\}\text{ and }\mathcal{S}_B=\{h_B(\eta_j),\ j=0,\dots,n-1\} \tag{9}$$

be the sets of the sampled values of h_A and h_B in the nodes θ_k and η_j given in (7). By the Grenander and Szegö theorem, for $n\to\infty$ the eigenvalues of H are equally distributed as the elements of the Cartesian product of $\mathcal{S}_A$ and $\mathcal{S}_B$.

3 The 1D fit preconditioner

We recall briefly the fit preconditioner, introduced in [6] for the 1D case. Hence in this section H is a 1D band Toeplitz matrix of size n, having the symmetric symbol function

$$h(\theta) = \sum_{j=-w}^{w} m_j \, \mathrm{e}^{\boldsymbol{i} j\theta} = m_0 + 2\sum_{j=1}^{w} m_j \, \cos j\theta, \quad \theta \in [-\pi, \pi].$$

By the Grenander and Szegö theorem, for $n \to \infty$ the eigenvalues of H are equally distributed as $h(\theta_k)$, with $\theta_k = -\pi + k\, 2\pi/n$, $k = 0, \ldots, n-1$. It follows that for a fixed n the set of the values $h(\theta_k)$ can be assumed as an acceptable approximation of the set of the eigenvalues of H.

The construction of the fit preconditioner is composed of two phases: the first one, called *regularization phase*, constructs a band Toeplitz positive definite matrix T whose greatest eigenvalues approximate the greatest eigenvalues of H, and the second one, called *factorization phase*, constructs a band lower triangular Toeplitz matrix L which approximates the Choleski factor of T.

3.1 Regularization phase for a positive definite H

If H is positive definite, consider the *cut function*

$$\varphi(\theta) = \max \, \{h(\theta), \tau\}. \tag{10}$$

The fit preconditioner is based on a trigonometric polynomial $t(\theta)$ of a suitable degree μ which satisfies the following requirements

- $t(\theta)$ is a good approximation of $\varphi(\theta)$,
- $t(\theta) \geq \rho\, \tau$ for a constant $0 < \rho < 1$.

Moreover $t(\theta)^{-1} h(\theta)$ should not oscillate too much. This can be obtained by choosing a small degree μ. Denoting by $\boldsymbol{\beta}(\theta) = [1, \cos\theta, \ldots, \cos\mu\theta]^T$ the cosine basis and by $\boldsymbol{c} = [c_0, \ldots, c_\mu]^T$ the $(\mu+1)$-vector of the coefficients, the polynomial we are looking for has the form $t(\theta) = \boldsymbol{c}^T \boldsymbol{\beta}(\theta)$.

There are many ways to find such a polynomial. We suggest the following one: let $\nu > 2\mu + 1$ be an integer and consider the set

$$\omega = \{\theta_k = -\pi + k\, 2\pi/\nu, \ k = 0, \ldots, \nu\}. \tag{11}$$

We require that $t(\theta)$ is a minimum norm approximation of $\varphi(\theta)$ on ω. If the spectral norm is chosen, the vector $\boldsymbol{c}$ is the solution of a quadratic problem. Let R be the matrix whose entries are

$$r_{k,j} = \beta_j(\theta_k), \quad \theta_k \in \omega, \quad \text{for } k = 0, \ldots, \nu, \ j = 0, \ldots, \mu,$$

and $\boldsymbol{\phi}$ be the vector whose entries are

$$\phi_k = \varphi(\theta_k), \quad \theta_k \in \omega, \quad \text{for } k = 0, \ldots, \nu.$$

Then $\boldsymbol{c}$ is the solution of the following quadratic problem

$$\|R\,\boldsymbol{c}-\boldsymbol{\phi}\|_2=\min_{\boldsymbol{p}}\|R\,\boldsymbol{p}-\boldsymbol{\phi}\|_2. \tag{12}$$

A set ω_1 of points not belonging to ω equally distributed in $[-\pi,\pi]$ is then chosen. If $t(\theta)<\rho\tau$ for $\theta\in\omega_1$, we can try a makeshift solution, obtained for example by choosing a greater value of ν, or by solving a suitable constrained quadratic problem. The one-dimensional fit polynomial is thus

$$t(\theta)=\sum_{j=0}^{\mu}c_j\cos j\theta. \tag{13}$$

Let T be the Toeplitz matrix generated by $t(\theta)$. If $t(\theta)$ is well chosen, the matrix $T^{-1}H$ has the greatest eigenvalues clustered around 1 and the smallest eigenvalues close to the corresponding eigenvalues of H divided by τ. If τ is not too small, this means that the smallest eigenvalues of $T^{-1}H$ are clustered around 0 [17].

3.2 Regularization phase for a non positive definite H

The fit preconditioner can be constructed also if H is not positive definite. The requirements on $t(\theta)$ guarantee that T is positive definite. Hence only the greatest positive eigenvalues of the preconditioned matrix are clustered around 1. If $h(\theta)\geq-\tau$, the preconditioned matrix has the negative eigenvalues in the cluster around zero. But if $h(\theta)<-\tau$, the preconditioned matrix could maintain some negative unclustered eigenvalues. This would have a deleterious effect on the convergence rate. In this case, instead of (10), the cut function

$$\varphi(\theta)=\max\ \left\{|h(\theta)|,\tau\right\} \tag{14}$$

should be considered. This position is analogous to that suggested in [11]. In this way the negative eigenvalues of the preconditioned matrix with the greatest modulus are clustered around -1. The polynomial $t(\theta)$ is then constructed as in the previous section.

3.3 Factorization phase

Matrix T is a band Toeplitz matrix with bandwidth μ. Its use as a preconditioner for CG requires solving at each iteration a linear system with matrix T. In order to reduce the cost, a factorization of T, which is definite positive, is suggested. The Choleski factor of T is not a Toeplitz matrix and would require more memory for its storage. Instead of it, a Toeplitz lower triangular matrix L is computed such that LL^T is close to T. This is easily obtained by computing the Wiener-Hopf factorization of the function $t(\theta)$ as follows: consider the Laurent polynomial

$$t(z)=c_0+\frac{1}{2}\sum_{j=1}^{\mu}c_j\,(z^j+z^{-j}),$$

whose restriction to the unit circle of the complex plane is the polynomial given in (13) (for the sake of simplicity we denote here and hereafter with the same name a function in the complex plane and its restriction on the unit circle), and find the function

$$\ell(z) = \sum_{j=-\mu}^{0} \ell_j z^j$$

such that $\ell(z^{-1}) \neq 0$ for $|z| \leq 1$, $\ell_0 > 0$ and $\ell(z)\ell(z^{-1}) = t(z)$. The coefficients ℓ_j, which satisfy the nonlinear system

$$\sum_{j=-\mu}^{0} \ell_j^2 = c_0 \quad \text{and} \quad \sum_{j=-\mu}^{k} \ell_j \, \ell_{j-k} = c_k/2, \quad k = -\mu, \ldots, -1, \tag{15}$$

can be computed by any of the methods described in [1, 4, 5, 15, 20]. The function

$$\ell(\theta) = \sum_{j=-\mu}^{0} \ell_j \, \mathrm{e}^{\boldsymbol{i} j\theta},$$

which will be called the *triangular factor* of $t(\theta)$, is the symbol function of L. The preconditioner is then $P = LL^T$. In [6] it is shown that no more than 2μ eigenvalues of the preconditioned matrix $P^{-1}H$ lie outside the set of the eigenvalues of the matrix generated by $t^{-1}(\theta)h(\theta)$.

4 The 2D fit preconditioner for a separable matrix

The construction of a fit preconditioner for the 2D case takes into account the separability of H and, for the nonseparable case, the symmetry level of the mask. In any case, the resulting fit preconditioner P is positive definite and has the form $P = LL^T$, where L is a lower triangular matrix with a 2D band Toeplitz structure. Hence its use has a cost of $O(n^2)$ operations per iteration. We examine in details the various cases.

When the matrix H is separable as in (3), its mask is a rank one matrix and its symbol is equal to the product of two functions $h_A(\theta)$ and $h_B(\eta)$, as in (8). The 2D fit preconditioner can be obtained by making the tensor product of two 1D fit preconditioners, constructed separately from $h_A(\theta)$ and $h_B(\eta)$ (this is the strategy proposed in [6]). Two (possibly different) regularization parameters τ_A and τ_B must be detected. Denote by n_A and n_B the cardinalities of the subsets of $\mathcal{S}_A$ and $\mathcal{S}_B$, whose elements are greater than τ_A and τ_B, respectively. Then τ_A and τ_B must be chosen in such a way that the product $n_A n_B$ is approximately equal to the cardinality of the signal subspace. Let $\varphi_A(\theta)$ and $\varphi_B(\eta)$ be the cut functions corresponding to $h_A(\theta)$ and $h_B(\eta)$ with the regularization parameters τ_A and τ_B respectively. Following section 3 the one dimensional fits $t_A(\theta)$ and $t_B(\eta)$ are found and from them the two triangular factors $\ell_A(\theta)$ and $\ell_B(\eta)$, which generate the two lower triangular matrices L_A and L_B, are computed. Then the matrix

$$P = L_A L_A^T \otimes L_B L_B^T = (L_A \otimes L_B)\,(L_A \otimes L_B)^T, \tag{16}$$

can be used as a preconditioner for the matrix H.

In addition to the technique described above, this separable case can be treated with the general techniques described in the following (Subsections 5.2 and 5.3) for the nonseparable case. As the experiments will show, the general techniques can be competitive.

5 The 2D fit preconditioner for a nonseparable matrix

When H is nonseparable, the 2D fit preconditioner can be constructed according to two different strategies:

- as the tensor product of two 1D fit preconditioners obtained by means of a separable approximation of H,
- from a single two-dimensional fit obtained directly from the mask of H.

5.1 Preconditioner constructed through a separable approximation of H

The standard and cheaper way to obtain a separable Toeplitz approximation K of H passes through a rank one approximation of the mask M, as described by the following result, derived from [14].

Let W be the $(2w+1)\times(2w+1)$ diagonal matrix whose principal elements are

$$w_{i,i} = \sqrt{n-|i|}, \quad \text{for} \quad i=-w,\dots,w.$$

For any pair of $(2w+1)$-vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, let A and B be the two $n\times n$ Toeplitz matrices defined in (4). Then

$$\|H - A\otimes B\|_F = \|WMW - (W\boldsymbol{a})(W\boldsymbol{b})^T\|_F.$$

Hence the best Toeplitz approximation in Frobenius norm among all separable 2D band Toeplitz matrices can be found by computing the singular value decomposition (SVD) of a $(2w+1)\times(2w+1)$ matrix. That is, let σ_1, $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ be the principal singular components of WMW. Then the matrix $\sigma_1\boldsymbol{u}_1\boldsymbol{v}_1^T$ is the best approximation of WMW in the spectral norm and in the Frobenius norm among all the rank one matrices. The vectors $\boldsymbol{a} = \sqrt{\sigma_1}\,W^{-1}\boldsymbol{u}_1$ and $\boldsymbol{b} = \sqrt{\sigma_1}\,W^{-1}\boldsymbol{v}_1$ are then obtained and from them the best separable approximation $K = A\otimes B$ of H.

An alternative way of computing a separable approximation of H uses directly the SVD of M. It leads to an optimal approximation of the symbol function $h(\theta,\eta)$ in the spectral norm. In fact, let $\boldsymbol{p}$ and $\boldsymbol{q}$ be two vectors of components p_i, q_i, $i=-w,\ldots,w$, with $p_{-i}=p_i$ and $q_{-i}=q_i$. Consider the two functions

$$p(\theta)=\sum_{k=-w}^{w} p_k\, \mathrm{e}^{\boldsymbol{i}k\theta} \quad \text{and} \quad q(\eta)=\sum_{j=-w}^{w} q_j\, \mathrm{e}^{\boldsymbol{i}j\eta},$$

for $\theta,\eta\in[-\pi,\pi]$. With the scalar product defined on $Q=[-\pi,\pi]^2$ by

$$< f,g >= \frac{1}{4\pi^2}\int_Q f(\theta,\eta)g(\theta,\eta)\ d\theta\ d\eta,$$

we have

$$\begin{aligned}\|h(\theta,\eta)-p(\theta)q(\eta)\|_2^2 &= \frac{1}{4\pi^2}\int_Q \left[h(\theta,\eta)-p(\theta)q(\eta)\right]^2\ d\theta\ d\eta \\ &= \sum_{k,j=-w}^{w}(m_{k,j}-p_kq_j)^2=\|M-\boldsymbol{p}\boldsymbol{q}^T\|_F^2.\end{aligned}$$

It follows that the functions $h_A(\theta)$ and $h_B(\eta)$ which solve the problem

$$\|h(\theta,\eta)-h_A(\theta)h_B(\eta)\|_2=\min_{p,q}\|h(\theta,\eta)-p(\theta)q(\eta)\|_2$$

are those having as coefficients the vectors $\boldsymbol{a}=\sqrt{\sigma_1}\boldsymbol{u}_1$ and $\boldsymbol{b}=\sqrt{\sigma_1}\boldsymbol{v}_1$, where σ_1, $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ are the principal singular components of M. The best separable approximation $K=A\otimes B$ of H is then obtained from $\boldsymbol{a}$ and $\boldsymbol{b}$.

The matrices A and B computed by the SVD of M or computed by the SVD of WMW are very close (in our experimentation the differences turn out to be negligible). Then we use simply the SVD of M in order to obtain the separable approximation K of H. The preconditioner (16) is then constructed from K, as described in Section 4.

5.2 Preconditioner obtained from a two-dimensional fit: regularization phase

The strategy of Subsection 5.1 is simple but effective only when K is sufficiently close to H, that is when the second singular value σ_2 of M is much smaller than the first one σ_1. When this is not true, the results can be poor. In such a case it is better to compute a two-dimensional fit directly from M. As stated in section 2, the set $\mathcal{S}_H$ can be assumed as an acceptable approximation of the set of the eigenvalues of H. In order to find the regularizing preconditioner based on a polynomial fit, consider the cut function

$$\varphi(\theta,\eta)=\max\ \left\{h(\theta,\eta),\tau\right\}, \tag{17}$$

if $h(\theta,\eta)\geq-\tau$. Otherwise consider the cut function

$$\varphi(\theta,\eta)=\max\ \left\{|h(\theta,\eta)|,\tau\right\}. \tag{18}$$

As in 1D case, the polynomial fit can be computed by solving a minimum constrained problem. We look for a trigonometric polynomial $t(\theta, \eta)$ which satisfies the following requirements

- $t(\theta, \eta)$ is a good approximation of $\varphi(\theta, \eta)$,
- $t(\theta, \eta) \geq \rho\, \tau$ for a constant $0 < \rho < 1$.

Moreover $t(\theta, \eta)^{-1} h(\theta, \eta)$ should not oscillate too much. This can be obtained by choosing a low degree for the polynomial.

Let $t(\theta, \eta) = \boldsymbol{c}^T \boldsymbol{\gamma}(\theta, \eta)$, where $\boldsymbol{c}$ is the vector of the coefficients of $t(\theta, \eta)$ in the basis $\boldsymbol{\gamma}(\theta, \eta)$. The computation of $\boldsymbol{c}$ can be made in a way similar to that described in section 3.1, $\boldsymbol{c}$ being the solution of a constrained quadratic problem of the form (12). We examine here two choices for the basis $\boldsymbol{\gamma}(\theta, \eta)$.

(a) *The 2-level symmetric case*: if the mask M is 2-level symmetric, the basis used for the fit is the $(\mu+1)^2$-vector $\boldsymbol{\gamma}(\theta, \eta)$ obtained by arranging rowwise the matrix $\boldsymbol{\beta}(\theta)\boldsymbol{\beta}^T(\eta)$, where $\boldsymbol{\beta}(\theta) = [1, \cos\theta, \ldots, \cos\mu\theta]^T$. The Cartesian products Ω of ω (defined in (11)) by itself can be chosen as the set of nodes on which evaluate the functions of the basis in order to construct the matrix R and the vector $\boldsymbol{\phi}$ as in Section 3.1. Vector $\boldsymbol{c}$ is the solution of the problem

$$\|R\,\boldsymbol{c} - \boldsymbol{\phi}\|_2 = \min_{\boldsymbol{p}} \|R\,\boldsymbol{p} - \boldsymbol{\phi}\|_2 \tag{19}$$

of size $(\nu+1)^2 \times (\mu+1)^2$. The function $t(\theta, \eta)$ is given by

$$t(\theta, \eta) = \boldsymbol{c}^T \gamma(\theta, \eta) = \sum_{k,j=0}^{\mu} c_{(\mu+1)k+j} \cos k\theta \cos j\eta.$$

Let N be the corresponding mask, which results to be 2-level symmetric.

(b) *The 1-level symmetric case*: even when the mask M is only 1-level symmetric, we can look for $t(\theta, \eta)$ expressed in terms of the same basis $\boldsymbol{\gamma}(\theta, \eta)$ as before, obtaining a Toeplitz positive definite matrix with a 2-level symmetric mask. This choice turns out to be effective if the asymmetry parameter α_M is low. For greater values of α_M the basis should include also the sine terms. More precisely, we complete $\boldsymbol{\gamma}(\theta, \eta)$ with the μ^2-vector obtained by arranging rowwise the matrix $\boldsymbol{\zeta}(\theta)\boldsymbol{\zeta}^T(\eta)$, where $\boldsymbol{\zeta}(\theta) = [\sin\theta, \ldots, \sin\mu\theta]^T$. The quadratic minimum problem to be solved in this case has size $(\nu+1)^2 \times [(\mu+1)^2 + \mu^2]$ and the form

$$\|\left[R \mid S\right] \overline{\boldsymbol{c}} - \boldsymbol{\phi}\|_2 = \min_{\boldsymbol{q}} \|\left[R \mid S\right] \boldsymbol{q} - \boldsymbol{\phi}\|_2. \tag{20}$$

Due to the ortogonality properties of the basis, the normal equations associated to (20) are

$$\left[R \mid S\right]^T \left[R \mid S\right] \overline{\boldsymbol{c}} - \left[R \mid S\right]^T \boldsymbol{\phi} = \begin{bmatrix} R^T R & O \\ O & S^T S \end{bmatrix} \overline{\boldsymbol{c}} - \begin{bmatrix} R^T \boldsymbol{\phi} \\ S^T \boldsymbol{\phi} \end{bmatrix} = \mathbf{0}.$$

Since $\nu > 2\mu + 1$, block $S^T S$ is nonsingular and has a scalar form, namely $S^T S = \nu^2/4\, I$. Then the solution $\overline{\boldsymbol{c}}$ is given by

$$\overline{\boldsymbol{c}} = \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{d} \end{bmatrix},$$

$\boldsymbol{c}$ being the solution of (19) and $\boldsymbol{d} = 4S^T \boldsymbol{\phi}/\nu^2$. In this case the function $t(\theta, \eta)$ is given by

$$t(\theta, \eta) = \sum_{k,j=0}^{\mu} c_{(\mu+1)k+j} \cos k\theta \cos j\eta + \sum_{k,j=1}^{\mu} d_{(k-1)\mu+j-1} \sin k\theta \sin j\eta$$

and the mask N results to be only 1-level symmetric.

5.3 Preconditioner obtained from a two-dimensional fit: factorization phase

Once $t(\theta, \eta)$ has been obtained, its use for constructing the preconditioner requires the computation of two suitable triangular factors. Generalizing the 1D case, we look for a factorization $g(z, v)g(z^{-1}, v^{-1})$ of $t(z, v)$. Let $t_{k,j}$, $k, j = -\mu, \ldots, \mu$, be the elements of the mask N and consider the Laurent polynomial

$$t(z, v) = \sum_{k,j=-\mu}^{\mu} t_{k,j}\, z^k\, v^j, \tag{21}$$

whose restriction to the unit circle of the complex plane is $t(\theta, \eta)$. Denoting by

$$\mathcal{I}_- = \big\{(k, j), \text{ with } \begin{array}{l} k = -\mu, \ldots, -1 \text{ and } j = -\mu, \ldots, \mu, \\ k = 0 \text{ and } j = -\mu, \ldots, 0 \ \big\}, \end{array} \tag{22}$$

a function

$$g(z, v) = \sum_{(k,j)\in\mathcal{I}_-} g_{k,j}\, z^k\, v^j, \tag{23}$$

such that $g(z^{-1}, v^{-1}) \neq 0$, for $|z| \leq 1$ and $|v| = 1$ and for $z = 0$ and $|v| \leq 1$ (this condition is generally referred to as the *stability* condition), and $g_{0,0} > 0$ would be acceptable provided that $g(z, v)\, g(z^{-1}v^{-1})$ is close to $t(z, v)$.

Unlike the 1D case, the Wiener-Hopf factorization of $t(z, v)$ into two factors of degree μ may not exist. In fact, if such factorization would exist, on the unit circle it would satisfy $t(\theta, \eta) = g(\theta, \eta)g(-\theta, -\eta) = |g(\theta, \eta)|^2$, but it is well known that a positive trigonometric polynomial of degree μ in the two variables θ and η cannot always be written as a sum of squares of polynomials in θ and η of the same degree [18]. Multilevel Toeplitz matrices can be seen as special cases of multi-index block Toeplitz matrices, whose spectral factorization is analyzed in [16]: under suitable hypotheses a spectral factorization of a positive two-variable symbol function $t(\theta, \eta)$ into two triangular factors belonging to the Wiener class exists, but even if $t(\theta, \eta)$ has a finite degree, the factors in general have an infinite expansion in one variable. Moreover the computation of these factors presents nontrivial difficulties from the numerical point of view. Since we look for a finite number of coefficients, an approximate factorization must be taken into account. It may therefore be convenient to consider also the factorization of the symbol $t^{-1}(\theta,\eta)$, which presents less numerical difficulties. The existence of such a factorization is shown in [16], under hypotheses which are satisfied in our case. We must note that also the factors obtained in this way have in general an infinite expansion (see [7]), which should be truncated to meet our aim.

In practice, we propose three different approximate factorizations: the first and the second ones are obtained from $t(\theta,\eta)$ and lead to a direct preconditioner, the third one is obtained from $t^{-1}(\theta,\eta)$ and leads to an inverse preconditioner.

5.3.1 Factorization by a rank one approximation of the mask N

The first factorization uses a rank one approximation of the mask N associated to $t(\theta,\eta)$. Let $\boldsymbol{a}=\sqrt{\sigma_1}\boldsymbol{u}_1$ and $\boldsymbol{b}=\sqrt{\sigma_1}\boldsymbol{v}_1$, where σ_1, $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ are the first singular components of N. The $2\mu+1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ entries are a_j and b_j, $j=-\mu,\ldots,\mu$ and verify $a_{-j}=a_j$ and $b_{-j}=b_j$. The following theorem guarantees that, if $t(\theta,\eta)>0$, the two trigonometric polynomials of degree μ

$$t_A(\theta)=a_0+2\sum_{j=1}^{\mu}a_j\,\cos j\theta,\qquad t_B(\eta)=b_0+2\sum_{j=1}^{\mu}b_j\,\cos j\eta,\tag{24}$$

are positive.

Theorem 1 *Let N be a 1-level symmetric mask defining the symbol function $t(\theta,\eta)$. If $t(\theta,\eta)$ is positive, the approximation $t_A(\theta)\,t_B(\eta)$ obtained through the principal singular components of N is also positive.*

Proof. Let ν be an integer, with $\nu\geq 2\mu+1$. Consider the following $\nu\times(2\mu+1)$ Fourier matrix Ω, whose (k,j)-th entry is $\nu^{-1/2}\ \mathrm{e}^{\boldsymbol{i}kj2\pi/\nu}$, for $k=0,\ldots,\nu-1$ and $j=-\mu,\ldots,\mu$. It is easy to verify that $\Omega^H\Omega=I$. Since $N^TN\boldsymbol{b}=\sigma_1^2\boldsymbol{b}$, we have

$$\widehat{N}^T\widehat{N}\,\widehat{\boldsymbol{b}}=\sigma_1^2\,\widehat{\boldsymbol{b}},\quad\text{where}\quad \widehat{N}=\Omega N\Omega^H,\quad \widehat{\boldsymbol{b}}=\Omega\,\boldsymbol{b}.$$

The (r,s)-th entry of $\widehat{N}$ and the r-th entry of $\widehat{\boldsymbol{b}}$ for $r,s=0,\ldots,\nu-1$ are

$$\widehat{m}_{r,s}=\frac{1}{\nu}\sum_{k,j=-\mu}^{\mu}m_{k,j}\ \mathrm{e}^{\boldsymbol{i}(kr-js)2\pi/\nu}=\frac{1}{\nu}\,t(\theta_r,-\eta_s),$$

$$\widehat{b}_r=\frac{1}{\sqrt{\nu}}\sum_{j=-\mu}^{\mu}b_j\ \mathrm{e}^{\boldsymbol{i}jr2\pi/\nu}=\frac{1}{\sqrt{\nu}}\,t_B(\eta_r),\quad\text{with}\quad \theta_r=\eta_r=r\,\frac{2\pi}{\nu}.$$

Then $\widehat{N}^T\widehat{N}\ >O$ and by Perron-Frobenius theorem the vector $\widehat{\boldsymbol{b}}$ can be chosen in such a way that $\widehat{\boldsymbol{b}}>\mathbf{0}$ and $\|\widehat{\boldsymbol{b}}\|_2=\sqrt{\sigma_1}$. Hence $\boldsymbol{b}$ can be chosen in such a way that $t_B(\eta_r) > 0$ for $r = 0, \ldots, \nu - 1$. From the arbitrariness of ν it follows that $t_B(\eta) > 0$ for any $\eta \in [-\pi, \pi]$. The proof for $t_A(\theta)$ is analogous. $\Box$

The positiveness of $t_A(\theta)$ and $t_B(\eta)$ allows us to compute the triangular factors $\ell_A(\theta)$ and $\ell_B(\eta)$, as described in Subsection 3.3. In this way we obtain the two lower triangular matrices L_A and L_B and from them the preconditioner (16).

It is worth noting that the factorization by a rank one approximation of the mask, applied to both the mask N obtained by any of the two methods of Subsection 5.2, gives the same result if the asymmetry of the 1-level symmetric mask is low, as shown in the following theorem.

Theorem 2 *Let N_a be the 2-level symmetric mask obtained in Section 5.2 (a) and N_b be the 1-level symmetric mask obtained in Section 5.2 (b). Let α_{N_b} be the asymmetry parameter of N_b and σ_1 be the first singular value of N_a. If $\alpha_{N_b} < 2\sigma_1$, then N_a and N_b have the same first singular components.*

Proof. Let J be the $\mu \times \mu$ exchange matrix. According to (5) and (6) we have

$$N_a = \begin{bmatrix} U & \boldsymbol{u} & UJ \\ \boldsymbol{v}^T & c_0 & \boldsymbol{v}^T J \\ JU & J\boldsymbol{u} & JUJ \end{bmatrix}, \quad N_b = N_a + \begin{bmatrix} V & \mathbf{0} & -VJ \\ \mathbf{0}^T & 0 & \mathbf{0}^T \\ -JV & \mathbf{0} & JVJ \end{bmatrix},$$

for suitable μ vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ and $\mu \times \mu$ matrices U and V, easily obtained from the solution of (20). Matrix N_a has at most $\mu + 1$ non zero singular values $\sigma_1, \ldots, \sigma_{\mu+1}$. Direct computations show that the singular values of N_b are $\sigma_1, \ldots, \sigma_{\mu+1}$ and the square roots of the eigenvalues of $4V^TV$. Hence, the first singular value of N_b is equal to $\max(\sigma_1, 2\|V\|_2)$. Since $\alpha_{N_b} = 4\|V\|_2$, if $\alpha_{N_b} < 2\sigma_1$, then σ_1 is the first singular value of N_b. Moreover the two matrices N_a and N_b have the same singular vectors corresponding to the singular values $\sigma_1, \ldots, \sigma_{\mu+1}$. $\Box$

In the numerical experimentation the hypothesis of the previous theorem is always verified.

5.3.2 Factorization of the function $t(\theta, \eta)$

The use of the approximate factorization of $t(\theta, \eta)$ into $t_A(\theta)t_B(\eta)$ can be effective only when the second singular value σ_2 of N is much smaller than the first one σ_1. This situation seldom arises when M does not have the first singular value much greater than the second one, since N is obtained through a two-dimensional fit starting from M. For this reason we suggest the following technique, which applies the Wiener-Hopf factorization to a 1D Laurent polynomial whose coefficients are obtained by flattening N.

Let $\boldsymbol{c}$ be the vector of entries $c_{(2\mu+1)k+j} = t_{k,j}$ for $k, j = -\mu, \ldots, \mu$. The function

$$\widetilde{t}(\eta) = \sum_{p=-\nu}^{\nu} c_p \, \mathrm{e}^{\boldsymbol{i} p\eta} = c_0 + 2\sum_{p=1}^{\nu} c_p \cos p\eta, \quad \text{where} \quad \nu = 2\mu^2 + 2\mu,$$

is the symbol function of a symmetric 1D Toeplitz matrix H_1, obtained from H by deleting the zero diagonals and filling the nonzero diagonals. Since $\widetilde{t}(\eta) = t\big((2\mu+1)\eta, \eta\big)$, then $\widetilde{t}(\eta)$ is positive for any η and its Wiener-Hopf factorization exists. Let

$$\ell(\eta) = \sum_{p=-\nu}^{0} \ell_p \, \mathrm{e}^{\boldsymbol{i} p\eta}$$

be the triangular factor of $\widetilde{t}(\eta)$. The coefficients ℓ_p, $p = -\nu, \ldots, 0$, satisfy system (15) and can be computed by any of the methods described in [1, 4, 5, 15, 20] (see subsection 3.3).

Consider now the $(2\mu+1) \times (2\mu+1)$ mask G whose elements are

$$g_{k,j} = \begin{cases} \ell_{(2\mu+1)k+j} & \text{for } (k,j) \in \mathcal{I}_-, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\ell(\eta) = \sum_{p=-\nu}^{0} \ell_p \, \mathrm{e}^{\boldsymbol{i} p\eta} = \sum_{(k,j)\in\mathcal{I}_-} g_{k,j} \, \mathrm{e}^{\boldsymbol{i}((2\mu+1)k+j)\eta}.$$

From the mask G we obtain the function defined in (23). Now we investigate when $g(z^{-1}, v^{-1})$ satisfies the stability condition. It is easy to show that this happens under hypotheses often verified in practice.

Theorem 3 *Let the mask N define the symbol function $t(\theta,\eta)$ positive for any θ,η. Let G be the mask obtained by applying Wiener-Hopf factorization to the flattened 1D mask, as previously described. If N is centerdominant, that is* $t_{0,0} > \sum\limits_{\substack{k,j=-\mu \\ (k,j)\neq(0,0)}}^{\mu} |t_{k,j}|$, *then*

(a) $\ell_0 > \sum\limits_{p=-\nu}^{-1} |\ell_p|$,

(b) $g(z^{-1}, v^{-1}) \neq 0$ *for* $|z| \leq 1$ *and* $|v| = 1$ *and for* $z = 0$ *and* $|v| \leq 1$.

Proof. The coefficients ℓ_p satisfy the relations

$$c_k = \sum_{j=-\nu}^{k} \ell_j \ell_{j-k} \quad \text{for} \quad k = -\nu, \ldots, 0. \tag{25}$$

They can be computed by means of the Bauer method [2], which considers the sequence $\{B^{(n)}\}_n$ of the Toeplitz matrices of size n generated by $\widetilde{t}(\eta)$. All $B^{(n)}$ are definite positive, so that the Cholesky factor $L^{(n)}$ of $B^{(n)}$ exists. Bauer proved that for $n \to \infty$ the last row $\big[0, \ldots, 0, l^{(n)}_{n,n-\nu}, \ldots, l^{(n)}_{n,n}\big]$ of $L^{(n)}$ approaches the vector $\big[0, \ldots, 0, \ell_{-\nu}, \ldots, \ell_0\big]$. Since N is centerdominant, $B^{(n)}$ is diagonally dominant for any n, and $L^{(n)}$ results to be also diagonally dominant. Hence

$$l^{(n)}_{n,n} > \sum_{p=-\nu}^{-1} \big|l^{(n)}_{n,n+p}\big|.$$

Letting $n \to \infty$ we get

$$\ell_0 \geq \sum_{p=-\nu}^{-1} |\ell_p|. \tag{26}$$

Moreover, from (25) we get

$$\begin{aligned}\Big(\ell_0 - \sum_{p=-\nu}^{-1} |\ell_p|\Big)^2 &= \sum_{p=-\nu}^{0} \ell_p^2 - 2\Big(\sum_{p=-\nu}^{-1} \ell_0|\ell_p| - \sum_{p=-\nu+1}^{-1}\sum_{j=-\nu}^{p-1} |\ell_j \ell_{j-p}|\Big) \\ &\geq c_0 - \sum_{\substack{p=-\nu \\ p\neq 0}}^{\nu} |c_p| = t_{0,0} - \sum_{\substack{(k,j)\in\mathcal{I}_- \\ (k,j)\neq(0,0)}} |t_{k,j}| > 0.\end{aligned}$$

Then relation (26) must hold with the strict inequality sign and (a) follows. Now we have

$$g(0, v^{-1}) = \sum_{j=-\mu}^{0} g_{0,j} v^{-j} = \sum_{j=-\mu}^{0} \ell_j v^{-j}.$$

Then for any v with $|v| \leq 1$ it results

$$|g(0, v^{-1})| \geq \Big|\ell_0 - \Big|\sum_{p=-\mu}^{-1} \ell_p v^{-p}\Big|\Big| \geq \ell_0 - \sum_{p=-\mu}^{-1} |\ell_p| > 0.$$

Moreover, for any z, v with $|z| \leq 1$ and $|v| = 1$ it results

$$\begin{aligned}|g(z^{-1}, v^{-1})| &= \Big|\sum_{(k,j)\in\mathcal{I}_-} g_{k,j} z^{-k} v^{-j}\Big| \geq \Big|\ell_0 - \Big|\sum_{\substack{(k,j)\in\mathcal{I}_- \\ k,j\neq(0,0)}} |\ell_{(2\mu+1)k+j} z^{-k} v^{-j}\Big|\Big| \\ &\geq \ell_0 - \sum_{p=-\nu}^{-1} |\ell_p| > 0. \quad \square\end{aligned}$$

Mask G defines the symbol function

$$g(\theta, \eta) = \sum_{(k,j)\in\mathcal{I}_-} g_{k,j}\ \mathrm{e}^{\boldsymbol{i}(k\theta + j\eta)} \tag{27}$$

of a lower triangular 2D Toeplitz matrix L with a block band of bandwidth μ and blocks with bandwidth μ (only the principal blocks are lower triangular). Matrix L, having principal elements equal to $g_{0,0} = \ell_0$, is nonsingular. The preconditioner we propose is $P = LL^T$.

Now we analyze how good is the approximation (27) by studying the difference between the Laurent polynomials $t(z, v)$ and $g(z, v)\, g(z^{-1}, v^{-1})$. It is easy to see that on the curve $\mathcal{C} : z = v^{2\mu+1}$ it is $g(z, v) = g(v^{2\mu+1}, v) = \ell(v)$, then $g(z, v)\, g(z^{-1}, v^{-1}) = t(z, v)$. Out of $\mathcal{C}$, denoting by

$$d_{k,j} = \sum_{r,s=-\mu}^{\mu} g_{r,s} g_{r-k,s-j}, \text{ for } k = -\mu, \ldots, \mu,\ j = -2\mu, \ldots, 2\mu,$$

we have

$$g(z,v)g(z^{-1},v^{-1}) = \sum_{k=-\mu}^{\mu} \sum_{j=-2\mu}^{2\mu} d_{k,j}\, z^k v^j. \tag{28}$$

On the other hand

$$\ell(v)\ell(v^{-1}) = \sum_{k=-\mu}^{\mu} \sum_{j=-2\mu}^{2\mu} d_{k,j} v^{(2\mu+1)k+j} = \sum_{k=-\mu}^{\mu} \sum_{j=-\mu}^{\mu} d_{k,j} v^{(2\mu+1)k+j}$$
$$+ \sum_{k=-\mu}^{\mu-1} \sum_{j=1}^{\mu} d_{k+1,j-2\mu-1} v^{(2\mu+1)k+j} + \sum_{k=-\mu+1}^{\mu} \sum_{j=-\mu}^{-1} d_{k-1,j+2\mu+1} v^{(2\mu+1)k+j}.$$

Since we required that $\widetilde{t}(v) = \ell(v)\ell(v^{-1})$, the coefficients of (21) result to be

$$t_{k,j} = \begin{cases} d_{k,j} + d_{k-1,j+2\mu+1} & \text{if } j = -\mu,\ldots,-1 \text{ and } k = -\mu+1,\ldots,\mu, \\ d_{k,j} + d_{k+1,j-2\mu-1} & \text{if } j = 1,\ldots,\mu \text{ and } k = -\mu,\ldots,\mu-1, \\ d_{k,j} & \text{otherwise.} \end{cases} \tag{29}$$

Then denoting by

$$\delta(z,v) = t(z,v) - g(z,v)\, g(z^{-1},v^{-1}) = \sum_{k=-\mu}^{\mu} \sum_{j=-2\mu}^{2\mu} \delta_{k,j}\, z^k v^j,$$

from (28) and (29) it follows that

$$\delta_{k,j} = -\delta_{k-1,2\mu+1+j} = \delta_{-k,-j} = -\delta_{-k+1,-2\mu-1-j} = -d_{k,j},$$
$$\text{for } k = -\mu+1,\ldots,\mu \text{ and } j = -2\mu,\ldots,-\mu-1.$$

The other coefficients are equal to zero. It can be shown that, for any k, in the sum defining $d_{k,j}$ the term $g_{k,j}g_{0,0}$ is present when $j = -\mu,\ldots,\mu$, while no term with the factor $g_{0,0}$ is present when $j = -2\mu,\ldots,-\mu-1$ or $j = \mu+1,\ldots,2\mu$. If the mask N is centerdominant, the mask G inherits this property, as seen in Theorem 3. Then the coefficients of $\delta(z,v)$ are small relatively to the coefficients of $t(z,v)$ and the symbol function $g(\theta,\eta)$ is an approximate factorization of $t(\theta,\eta)$ sufficiently good for our aim, as confirmed by the numerical experiments shown in Section 8.

5.3.3 Factorization of the function $t^{-1}(\theta,\eta)$

As previously told, a triangular factor of $t^{-1}(\theta,\eta)$ has in general an infinite expansion. Then we look for a finite approximation of it, that is a function of the form

$$q(\theta,\eta) = \sum_{(k,j)\in\mathcal{I}_-} q_{k,j}\, \mathrm{e}^{\boldsymbol{i}(k\theta+j\eta)}, \tag{30}$$

such that $q(z,v)q(z^{-1},v^{-1})$ and $t^{-1}(z,v)$ have the same coefficients of the terms $z^k v^j$ for $|k|,|j| \le \mu$. From (30) it follows that $q^{-1}(z^{-1},v^{-1})$ lacks the terms $z^k v^j$

with $(k,j) \in \mathcal{I}_- \backslash \{(0,0)\}$ and has a constant term equal to $q_{0,0}^{-1}$. By requiring the coefficient of the term $z^k v^j$ of $t(z,v)q(z,v)$ to be equal to the corresponding one of $q^{-1}(z^{-1},v^{-1})$ for any $(k,j) \in \mathcal{I}_-$, we obtain the $\mu' = 2\mu^2 + 2\mu + 1$ equations

$$\sum_{(r,s)\in\mathcal{I}_-} t_{k-r,j-s} q_{r,s} = \begin{cases} q_{0,0}^{-1} & \text{if } (k,j) = (0,0), \\ 0 & \text{otherwise,} \end{cases} \tag{31}$$

where $t_{k,j} = 0$ for indices outside the interval $[-\mu, \mu]$.

Let S be the leading principal minor of size μ' of the 2D Toeplitz matrix of size $(2\mu+1)^2$ generated by $t(\theta,\eta)$. Calling by Q the mask of $q(\theta,\eta)$, let $\boldsymbol{q}$ be the vector obtained from Q by ordering rowwise the elements $q_{k,j}$ for $k,j \in \mathcal{I}_-$, that is

$$\boldsymbol{q} = \left[q_{-\mu,-\mu}, q_{-\mu,-\mu+1}, \ldots, q_{0,-1}, q_{0,0}\right]^T.$$

From (31) it follows that $\boldsymbol{q}$ is the solution of the system

$$S\boldsymbol{q} = q_{0,0}^{-1} \boldsymbol{e}_{\mu'},$$

where $\boldsymbol{e}_{\mu'}$ is the last vector of the canonical base. Hence $q_{0,0}\boldsymbol{q}$ is the last column of the inverse of S. If the mask N is centerdominant, S is diagonally dominant and the last column of the inverse can be stably computed.

Let R be the 2D Toeplitz matrix associated to Q. R is lower triangular with a block band of bandwidth μ and blocks with bandwidth μ (only the principal blocks are lower triangular). The preconditioner we propose is $P = (RR^T)^{-1}$. It is applied in the inverse form, then the preconditioned matrix is $P^{-1}H = RR^TH$. Let $\nu \geq \mu$ be any integer. Denote by $S^{(\nu)}$ the leading principal minor of size $\nu' = 2\nu^2 + 2\nu + 1$ of the 2D Toeplitz matrix of size $(2\nu+1)^2$ generated by $t(\theta,\eta)$. Let $\boldsymbol{u}^{(\nu)}$ be the last column of the inverse of $S^{(\nu)}$ and set $\boldsymbol{q}^{(\nu)} = \boldsymbol{u}^{(\nu)}/\sqrt{u_{\nu'}}$, where $u_{\nu'}$ is the last element of $\boldsymbol{u}^{(\nu)}$. Then $\boldsymbol{q} = \boldsymbol{q}^{(\mu)}$.

Let now $\mathcal{I}_-^{(\nu)}$ be the set of the indices of $\boldsymbol{q}^{(\nu)}$. Let $\boldsymbol{q}^\infty$ be the infinite vector of the coefficients of the exact triangular factor of $t^{-1}(\theta,\eta)$ and $\boldsymbol{q}_\nu^\infty$ be the subset of $\boldsymbol{q}^\infty$ obtained by choosing the elements with indices in $\mathcal{I}_-^{(\nu)}$. In [16] it is proved that

$$\lim_{\nu\to\infty} \|\boldsymbol{q}_\nu^\infty - \boldsymbol{q}^{(\nu)}\| = 0$$

for any norm. Hence, for ν sufficiently large, $\boldsymbol{q}^{(\nu)}$ would be a good approximation of the coefficients of the exact triangular factor of $t^{-1}(\theta,\eta)$.

In order to analyze how good is the approximation we have obtained with $\boldsymbol{q}$, let $\nu > \mu$. The matrix $S^{(\mu)}$ is easily recognized as a principal submatrix of $S^{(\nu)}$, since a permutation matrix Π of size ν' exists such that

$$\widetilde{S}^{(\nu)} = \Pi S^{(\nu)} \Pi^T = \begin{bmatrix} S_1 & S_2 \\ S_2^T & S^{(\mu)} \end{bmatrix},$$

where S_1 and S_2 are blocks of size $(\nu' - \mu') \times (\nu' - \mu')$ and $(\nu' - \mu') \times \mu'$ respectively. Let $\boldsymbol{z}^{(\nu)}$ be the subvector of the last μ' components of $\Pi\boldsymbol{u}^{(\nu)}$, which is the last column of the inverse of $\widetilde{S}^{(\nu)}$. It is easy to show that

$$\boldsymbol{z}^{(\nu)} = \boldsymbol{u}^{(\mu)} + \boldsymbol{v}, \quad \text{with} \quad \boldsymbol{v} = {S^{(\mu)}}^{-1} S_2^T T^{-1} S_2 \boldsymbol{u}^{(\mu)},$$

where $T = S_1 - S_2 {S^{(\mu)}}^{-1} S_2^T$ is the Schur complement of $S^{(\mu)}$ in $\widetilde{S}^{(\nu)}$. The blocks $S^{(\mu)}$ and S_1 have a Toeplitz or nearly Toeplitz structure, with the diagonal elements equal to $t_{0,0}$, while $t_{0,0}$ is not present in S_2. If the mask N is centerdominant, then $\boldsymbol{v}$ is small compared with $\boldsymbol{u}^{(\mu)}$, that is $\boldsymbol{q}$ is a sufficiently good approximation of $\boldsymbol{q}^\infty$ for our aim, as confirmed by the numerical experiments shown in Section 8.

6 Eigenvalues of the preconditioned matrix

Starting from one of the two-dimensional fits $t(\theta,\eta)$ obtained in Section 5.2, in section 5.3 we have introduced three approximate factorizations, which can be regarded as the exact factorizations of a function $s(\theta,\eta)$ sufficiently close to $t(\theta,\eta)$. Namely on the unit circle we have

$$s(\theta,\eta) = \begin{cases} \ell_A(\theta)\ell_B(\eta)\ell_A(-\theta)\ell_B(-\eta) & \text{(see 5.3.1)}, \\ g(\theta,\eta)g(-\theta,-\eta) & \text{(see 5.3.2)}, \\ q^{-1}(\theta,\eta)q^{-1}(-\theta,-\eta) & \text{(see 5.3.3)}. \end{cases} \tag{32}$$

The difference $\delta(\theta,\eta) = t(\theta,\eta) - s(\theta,\eta)$ depends on problem features which are not easily evaluated a-priori. For the factorization 5.3.1, $\delta(\theta,\eta)$ is small when $t(\theta,\eta)$ is nearly separable; for the factorizations 5.3.2 and 5.3.3, $\delta(\theta,\eta)$ depends on the diagonal dominance of the 2D Toeplitz matrix associated to $t(\theta,\eta)$. In any case we assume that $\delta(\theta,\eta)$ is sufficiently small for our purpose, that is, we assume that the 2D Toeplitz matrix Z associated to the function $s^{-1}(\theta,\eta)h(\theta,\eta)$ has a suitable selected clustering of the largest eigenvalues around 1 with regularizing effect.

In the case of a symmetric definite positive coefficient matrix, the behaviour of PCG depends on the clustering of the spectrum of the preconditioned matrix. Thus, if we could use matrix Z as the preconditioned matrix, the assumption we have made would guarantee a quick convergence. But the preconditioner we propose is obtained through a factorization step, that is, the preconditioned matrix has the form $L^{-T}L^{-1}H$ for the direct preconditioner and RR^TH for the inverse preconditioner. On account of this factorization, some outliers greater than 1 occur in the spectrum of the preconditioned matrix. Many such outliers can reduce considerably the convergence speed. Hence we are interested in finding how many outliers the preconditioned matrix, corresponding to the three factorizations, has.

Lemma 1 *Let $\mathcal{I}_-$ be the set defined in (22). Let*

$$f(\theta,\eta) = \sum_{k,j=-\infty}^{\infty} f_{k,j}\ \mathrm{e}^{\boldsymbol{i}(k\theta+j\eta)}$$

and

$$c(\theta,\eta) = \sum_{(k,j)\in\mathcal{I}_-} c_{k,j}\ \mathrm{e}^{\boldsymbol{i}(k\theta+j\eta)}$$

be the symbol functions associated to the 2D $n \times n$ Toeplitz matrices F and C respectively. C is a lower triangular 2D Toeplitz matrix with a block band of bandwidth μ and blocks with bandwidth μ, except the principal blocks which are lower triangular. Let

$$y(\theta, \eta) = c(\theta, \eta) f(\theta, \eta) c(-\theta, -\eta).$$

be the symbol function of the 2D $n \times n$ Toeplitz matrix Y. Define $E = CFC^T$. Then the rank of the matrix $Y - E$ is bounded from above by $2(\mu n + 2(n-\mu)\mu)$.

Proof. The function $y(\theta, \eta)$ has the expansion

$$y(\theta, \eta) = \sum_{\omega, u=-\infty}^{\infty} y_{\omega,u} \exp(\boldsymbol{i}(\omega\theta + u\eta)),$$

where

$$y_{\omega,u} = \sum_{(\rho,v)\in\mathcal{I}_-} \sum_{(\sigma+\rho,z+v)\in\mathcal{I}_-} c_{\sigma+\rho,z+v}\, f_{\omega-\sigma,u-z}\, c_{\rho,v}. \tag{33}$$

Since Y is a 2D Toeplitz matrix, its (k,j)th element of the (α,β)th block is given by

$$y_{k,j}^{(\alpha,\beta)} = y_{\beta-\alpha,j-k}, \quad \text{for} \quad \alpha, \beta, k, j = 1, \dots, n.$$

The (k,j)th element of the (α,β)th block of E, with $\alpha, \beta, k, j = 1, \dots, n$, is equal to

$$e_{k,j}^{(\alpha,\beta)} = \sum_{\gamma=\gamma_1}^{\alpha} \sum_{\delta=\delta_1}^{\beta} \sum_{i=i_1}^{i_2} \sum_{s=s_1}^{s_2} c_{k,i}^{(\alpha,\gamma)}\, f_{i,s}^{(\gamma,\delta)}\, c_{j,s}^{(\beta,\delta)},$$

where

$$\gamma_1 = \max(\alpha-\mu, 1), \quad \delta_1 = \max(\beta-\mu, 1), \quad i_1 = \max(k-\mu, 1),$$
$$i_2 = \min(k+\mu, n), \quad s_1 = \max(j-\mu, 1), \quad s_2 = \min(j+\mu, n).$$

Actually, due to the lower triangular form of the principal blocks of C, we have $i_2 = k$ when $\gamma = \alpha$ and $s_2 = j$ when $\delta = \beta$, but this additional condition will not be stressed on, since it does not affect the result. By simple algebra we can show that

$$e_{k,j}^{(\alpha,\beta)} = \sum_{\rho=\rho_1}^{0} \sum_{\sigma+\rho=\sigma_1}^{0} \sum_{v=v_1}^{v_2} \sum_{z+v=z_1}^{z_2} c_{\sigma+\rho,z+v}\, f_{\beta-\alpha-\sigma,j-k-z}\, c_{\rho,v}, \tag{34}$$

where

$$\rho_1 = \max(-\mu, 1-\beta), \quad \sigma_1 = \max(-\mu, 1-\alpha), \quad v_1 = \max(-\mu, 1-j),$$
$$v_2 = \min(\mu, n-j), \quad z_1 = \max(-\mu, 1-k), \quad z_2 = \min(\mu, n-k).$$

By comparing relation (34) with relation (33) for $\omega = \beta - \alpha$ and $u = j - k$, we can see that matrices Y and E have the same elements with indices $(k,j), (\alpha,\beta)$ such that

$$\rho_1 = -\mu, \quad \sigma_1 = -\mu, \quad v_1 = -\mu, \quad v_2 = \mu, \quad z_1 = -\mu, \quad z_2 = \mu,$$

that is, when

$$\alpha \geq 1+\mu, \quad \beta \geq 1+\mu, \quad 1+\mu \leq j \leq n-\mu, \quad 1+\mu \leq k \leq n-\mu.$$

Hence nonzero elements of $Y-E$ can be found only in the first μ block rows and columns and in the first and last μ rows and columns of the remaining blocks. $\square$

The following analysis of the spectrum of the preconditioned matrices makes use of the Cauchy interlacing theorem which applies to symmetric matrices. Thus the symmetric similar matrix $L^{-1}HL^{-T}$ is considered instead of $L^{-T}L^{-1}H$ for the direct preconditioners, and the symmetric similar matrix R^THR is considered instead of RR^TH for the inverse preconditioner.

Theorem 4 *Let $h(\theta,\eta)$ be the symbol function of the 2D $n\times n$ Toeplitz matrix H of system (1) and $s(\theta,\eta)$ be the symbol function defined in (32) as a suitable approximation of the two-dimensional fit, according to one of the three factorizations 5.3.1, 5.3.2, 5.3.3. Let Z be the 2D $n\times n$ Toeplitz matrix associated to $s^{-1}(\theta,\eta)h(\theta,\eta)$. At most $O(n)$ eigenvalues of the preconditioned matrix lie outside the spectrum of Z.*

Proof. Since the function $s^{-1}(\theta,\eta)h(\theta,\eta)$ is real for $|\theta|=|\eta|=1$, the matrix Z is symmetric. We examine first the preconditioner obtained in 5.3.2. In this case we consider

$$Z-L^{-1}HL^{-T}=L^{-1}(LZL^T-H)L^{-T}.$$

From Lemma 1, applied to functions $f(\theta,\eta)=s^{-1}(\theta,\eta)h(\theta,\eta)$ and $c(\theta,\eta)=g(\theta,\eta)$, which generates L, it follows that the symmetric matrix LZL^T-H has rank $O(n)$. Then by the Cauchy interlacing theorem, no more than $O(n)$ eigenvalues of $L^{-1}HL^{-T}$ lie outside the cluster of the eigenvalues of Z.

We examine next the preconditioner obtained in 5.3.1. In this case we consider

$$Z-(L_A\otimes L_B)^{-1}H(L_A\otimes L_B)^{-T}.$$

The thesis follows as before, by applying Lemma 1 to functions $f(\theta,\eta)=s^{-1}(\theta,\eta)h(\theta,\eta)$ and $c(\theta,\eta)=\ell_A(\theta)\ell_B(\eta)$ which generates $L_A\otimes L_B$.

Finally we examine the preconditioner obtained in 5.3.3. In this case we should consider the matrix

$$Z-R^THR.$$

By observing that R^THR is similar to RHR^T, since $R^THR=J(RHR^T)J$ where J is the $n^2\times n^2$ exchange matrix, we apply Lemma 1 to functions $f(\theta,\eta)=h(\theta,\eta)$ and $c(\theta,\eta)=q(\theta,\eta)$ which generates R. By the Cauchy interlacing theorem, no more than $O(n)$ eigenvalues of RHR^T lie outside the cluster of the eigenvalues of Z. $\square$

In conclusion, all the factorizations we have proposed appear to be asymptotically equivalent from the point of view of the outliers number.

Figure 1 shows the spectra of two preconditioned matrices. The matrix H is associated to a 1-level symmetric mask with asymmetry parameter $\alpha_M = 0.033$. The preconditioners are obtained by computing a two-dimensional fit (as described in 5.2) with the 2-level symmetric basis (dotted line) and with the 1-level symmetric basis (continuous line), followed by the factorization of $t(\theta, \eta)$ described in 5.3.2. The eigenvalues of the nonpreconditioned matrix (gray line) are also shown. The presence of outliers is evident in both the spectra of the preconditioned matrices. The fit obtained by using the 1-level symmetric basis appears to be more suitable, since the largest eigenvalues of the preconditioned matrix are more clustered around 1.

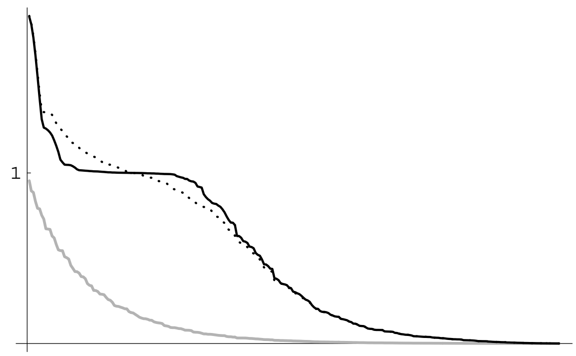

Figure 1: Eigenvalues of the preconditioned.

7 Computational costs

The cost of each iteration of CG or MRII is obtained by adding the cost of computing $H\boldsymbol{z}$ for $\boldsymbol{z} \in \mathbb{R}^{n^2}$ and the cost of computing $P^{-1}\boldsymbol{z}$ for $\boldsymbol{z} \in \mathbb{R}^{n^2}$. This last computation is actually performed by solving linear systems with triangular Toeplitz matrices if the preconditioner is applied in the direct form and by performing products of Toeplitz matrices times vectors if the preconditioner is applied in the inverse form.

The computation of $H\boldsymbol{z}$ has cost

- $2wn^2$ if H is separable,
- $4w^2n^2$ if H is nonseparable.

The computation of $P^{-1}\boldsymbol{z}$ has cost

- $4\mu n^2$ with the separable preconditioner (16). Such a preconditioner is obtained either directly, as in Section 4, or through a rank one approximation as in Subsections 5.1 or 5.3.1,

- $4\mu^2 n^2$ with the nonseparable preconditioner, regardless of the fact that the preconditioner is applied in the direct form, as in Subsection 5.3.2, or in the inverse form, as in Subsection 5.3.3.

Thus the use of the nonseparable preconditioner justifies only if the number of iterations is sufficiently lower than the number of iterations of a separable preconditioner, especially when H is separable, as we will see in the next section.

8 Numerical experiments

In the previous section we have proposed and analyzed some regularizing preconditioners, based on fit techniques and approximate factorization strategies. In this section we perform a numerical experimentation in order to test the reconstruction efficiency of the proposed preconditioners. For comparison purpose the modified Chan preconditioner [12] is also considered with the optimal value τ_{ch} of the regularization parameter (detected, for each problem, through an ad-hoc experimentation).

8.1 Test problems

The experiments have been conducted on the 128×128 image shown in Figure 2 (synthetic Hoffman brain phantom [13]).

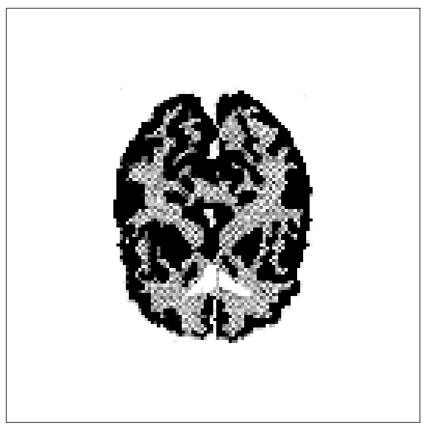

Figure 2: Original Hoffman phantom image.

The following masks M, depending on positive parameters α, β, γ and representing various PSFs, have been considered. In all cases the entries of M are scaled by the constant γ in such a way that $\sum_{i,j} m_{i,j} = 1$. The bandwidth is always $w = 8$.

- The mask of the Gaussian PSF is given by

$$m_{i,j} = \gamma \mathrm{e}^{-\alpha\, i^2 - \beta\, j^2}, \qquad i, j = -w, \ldots, w.$$

It has rank one, hence it is 2-level symmetric.

• The diffraction in incoherent illumination PSFs describes the diffraction effects caused by a system of lenses in a spatially incoherent illumination [3]. The mask for a circular pupil is given by

$$m_{i,j} = \gamma J_1^2(\alpha\sqrt{i^2+j^2})/(i^2+j^2), \qquad i,j=-w,\ldots,w,$$

where α depends on the radius of the pupil. It is 2-level symmetric.

• The mask of a motion PSF is given by

$$m_{i,j} = \gamma \mathrm{e}^{-\alpha\,(i+j)^2-\beta(i-j)^2}, \qquad i,j=-w,\ldots,w.$$

If $\alpha \neq \beta$, it is 1-level symmetric.

Various values of parameters α and β have been considered, obtaining masks with different properties.

The noisy image $\boldsymbol{f}$ is obtained by computing $H\boldsymbol{x}+\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a vector of randomly generated entries, with normal distribution and mean 0, scaled in such a way that $\|\boldsymbol{\eta}\|_2/\|H\boldsymbol{x}\|_2 = 10^{-4}$. The dimension of the noise and signal subspaces are estimated and reasonable values for the regularization parameter τ are found.

8.2 Numerical results

The problems with positive definite matrices have been solved by applying CG, while the problems with indefinite matrices have been solved by applying MRII. When CG is applied, the preconditioner is used in the left version, that is the preconditioned matrix is $L^{-T}L^{-1}H$ (in direct preconditioning) and RR^TH (in inverse preconditioning). When MRII is applied, the preconditioner is used in the split version, that is the preconditioned matrix is $L^{-1}HL^{-T}$ (in direct preconditioning) and R^THR (in inverse preconditioning).

Denote by $\boldsymbol{x}^{(i)}$ the vector obtained at the ith iteration starting with $\boldsymbol{x}^{(0)} = \mathbf{0}$ and by $e^{(i)} = \|\boldsymbol{x}^{(i)} - \boldsymbol{x}\|_2/\|\boldsymbol{x}\|_2$ the relative error.

For any mask, firstly we apply to the problem the non-preconditioned CG or MRII, in order to determine the reconstruction efficiency limit. That is, we consider the minimum error $e_m = \min_i e^{(i)}$; the quantity $E = 1.01\ e_m$ is taken as the reference value, in the sense that any approximated image having an error lower than E is considered as an acceptable reconstruction. Then the preconditioners described in Sections 4 and 5 are applied. They are obtained by combining the regularizing and factorizing techniques described above and are denoted in the following way:

- 1D fits denotes the preconditioner described in 4,
- rank one + 1D fits denotes the preconditioner described in 5.1,

- 2-level sym + rank one, 2-level sym + direct fact and 2-level sym + inverse fact denote the preconditioners obtained by constructing the two-dimensional fit on a basis formed by only cosine terms, as described in 5.2(a), followed by the factorization described in 5.3.1, 5.3.2 and 5.3.3, respectively,
- 1-level sym + direct fact and 1-level sym + inverse fact denote the preconditioners obtained by constructing the two-dimensional fit on a basis of sine and cosine terms, as described in 5.2(b), followed by the factorization described in 5.3.2 and 5.3.3, respectively.

The results are summarized in a table, where in the column denoted by "% error" is shown the notation "acc" and in the column denoted by "iterations" is shown the minimum iteration number κ such that $e^{(\kappa)} \leq E$, if an acceptable solution has been obtained. If an acceptable solution has not been obtained, in the second column the quantity $|\min_i e^{(i)} - e_m|/e_m$ and in the third column the index κ such that $e^{(\kappa)} = \min_i e^{(i)}$ are shown. At the end of the table the results of the modified Chan preconditioner are also shown.

8.2.1 Separable case

The experiments of the separable case have been conducted with the Gaussian mask with two choices of the parameters: with the first choice (see Table 1) matrix H is positive definite while with the second choice (see Table 2) matrix H is indefinite. We denote by τ_1 and τ_2 the regularization parameters used for the 1D fits strategy, by τ_3 the regularization parameter used for the two-dimensional fit strategies, and by τ_{ch} the regularization parameter used for the modified Chan preconditioner.

method	% error	iterations
Non-preconditioned CG	acc	20
1D fits	acc	11
2-level sym + direct fact	acc	6
2-level sym + inverse fact	acc	6
Chan preconditioner	acc	12

Table 1: Gaussian mask, $\alpha = 0.4$, $\beta = 0.2$, $e_m = 0.0703$, $\tau_1 = 0.5$, $\tau_2 = 0.4$, $\tau_3 = 0.03$, $\tau_{ch} = 0.15$.

As shown in Tables 1 and 2 the preconditioners based on a two-dimensional fit gives an acceptable reconstruction with fewer iterations than the 1D fits preconditioner obtained from the factors of H. A possible explanation of this behaviour lies in the fact that the former preconditioner allows a better separation between the noise eigenvalues and the cluster of the signal eigenvalues. However, if we take into account the cost of each iteration, the overall computational cost

method	% error	iterations
Non-preconditioned MRII	acc	72
1D fits	acc	34
2-level sym + direct fact	acc	28
2-level sym + inverse fact	acc	24
Chan preconditioner	acc	45

Table 2: Gaussian mask, $\alpha = 0.1$, $\beta = 0.1$, $e_m = 0.0973$, $\tau_1 = 0.5$, $\tau_2 = 0.5$, $\tau_3 = 0.06$, $\tau_{ch} = 0.25$.

of the 1D fits preconditioner is lower than the cost of the preconditioners based on a two-dimensional fit.

8.2.2 Nonseparable positive definite case

Hereafter we denote by τ_1 the regularization parameter used for the rank one + 1D fits strategy, by τ_2 the regularization parameter used for the two-dimensional fit strategies, and by τ_{ch} the regularization parameter used for the modified Chan preconditioner. A first experiment has been conducted with the circular pupil mask. The chosen value $\alpha = 2.2$ guarantees the positive definiteness of matrix H. In this case $\sigma_1/\sigma_2 = 18$, that is the mask is nearly separable. The results are shown in Table 3. A second set of experiments concerns the motion mask,

method	% error	iterations
Non-preconditioned CG	acc	41
Rank one + 1D fits	acc	15
2-level sym + rank one	acc	15
2-level sym + direct fact	acc	4
2-level sym + inverse fact	acc	7
Chan preconditioner	acc	5

Table 3: Circular pupil mask, $\alpha = 2.2$, $\sigma_1/\sigma_2 = 18$, $e_m = 0.0015$, $\tau_1 = 0.01$, $\tau_2 = 0.005$, $\tau_{ch} = 0.005$.

with values of the parameters α and β which guarantee the positive definiteness of matrix H. The different cases show an increasing asymmetry parameter and a decreasing ratio σ_1/σ_2. Tables 4, 5 and 6 summarize the results. For any case the ratio σ_1/σ_2 and the value of α_M are shown in the caption of the table.

8.2.3 Nonseparable indefinite case

A first experiment has been conducted with the circular pupil mask. The chosen value $\alpha = 1.1$ makes matrix H indefinite. In this case $\sigma_1/\sigma_2 = 31$, that is the

method	% error	iterations
Non-preconditioned CG	acc	19
Rank one + 1D fits	acc	12
2-level sym + rank one	1.5%	11
2-level sym + direct fact	acc	9
2-level sym + inverse fact	acc	9
1-level sym + direct fact	acc	9
1-level sym + inverse fact	acc	9
Chan preconditioner	acc	13

Table 4: Motion mask, $\alpha = 0.11$, $\beta = 0.1$, $\sigma_1/\sigma_2 = 42$, $\alpha_M = 0.0087$, $e_m = 0.0805$, $\tau_1 = 0.5$, $\tau_2 = 0.2$, $\tau_{ch} = 0.24$.

method	% error	iterations
Non-preconditioned CG	acc	21
Rank one + 1D fits	acc	15
2-level sym + rank one	1.5%	12
2-level sym + direct fact	acc	11
2-level sym + inverse fact	acc	11
1-level sym + direct fact	acc	9
1-level sym + inverse fact	acc	9
Chan preconditioner	acc	15

Table 5: Motion mask, $\alpha = 0.2$, $\beta = 0.05$, $\sigma_1/\sigma_2 = 3$, $\alpha_M = 0.11$, $e_m = 0.0825$, $\tau_1 = 0.4$, $\tau_2 = 0.15$, $\tau_{ch} = 0.2$.

mask is practically separable. The results are shown in Table 7. A second set of experiments concerns the motion mask, with values of the parameters α and β such that matrix H results to be indefinite. The different cases show an increasing asymmetry parameter and a decreasing ratio σ_1/σ_2. Tables 8, 9 and 10 summarize the results.

We observe that

- even when the ratio σ_1/σ_2 is high, corresponding to a nearly separable case, the techniques based on a rank one approximation likely require more iterations than the others. The lower the ratio, the more iterations are required.

- Applying rank one approximation to a two-dimensional fit does not appear to give better results than finding rank one approximation of the original mask and then making the regularization.

method	% error	iterations
Non-preconditioned CG	acc	25
Rank one + 1D fits	acc	23
2-level sym + rank one	acc	19
2-level sym + direct fact	acc	16
2-level sym + inverse fact	acc	16
1-level sym + direct fact	acc	12
1-level sym + inverse fact	acc	12
Chan preconditioner	acc	16

Table 6: Motion mask, $\alpha = 0.05$, $\beta = 0.7$, $\sigma_1/\sigma_2 = 1.73$, $\alpha_M = 0.23$, $e_m = 0.0619$, $\tau_1 = 0.55$, $\tau_2 = 0.2$, $\tau_{ch} = 0.2$.

method	% error	iterations
Non-preconditioned MRII	acc	46
Rank one + 1D fits	acc	20
2-level sym + rank one	acc	20
2-level sym + direct fact	acc	17
2-level sym + inverse fact	acc	17
Chan preconditioner	acc	29

Table 7: Circular pupil mask, $\alpha = 1.1$, $\sigma_1/\sigma_2 = 31$, $e_m = 0.0789$, $\tau_1 = 0.17$, $\tau_2 = 0.1$, $\tau_{ch} = 0.3$.

method	% error	iterations
Non-preconditioned MRII	acc	77
Rank one + 1D fits	acc	39
2-level sym + rank one	1.5%	32
2-level sym + direct fact	acc	33
2-level sym + inverse fact	acc	32
1-level sym + direct fact	acc	32
1-level sym + inverse fact	acc	32
Chan preconditioner	acc	46

Table 8: Motion mask, $\alpha = 0.04$, $\beta = 0.05$, $\sigma_1/\sigma_2 = 18$, $\alpha_M = 0.013$, $e_m = 0.1$, $\tau_1 = 0.35$, $\tau_2 = 0.1$, $\tau_{ch} = 0.15$.

- When α_M increases, that is the blocks of H become less and less symmetric, it appears to be convenient to use a 1-level symmetric fit strategy.

method	% error	iterations
Non-preconditioned MRII	acc	67
Rank one + 1D fits	acc	37
2-level sym + rank one	3.3%	18
2-level sym + direct fact	acc	22
2-level sym + inverse fact	acc	22
1-level sym + direct fact	acc	22
1-level sym + inverse fact	acc	22
Chan preconditioner	acc	41

Table 9: Motion mask, $\alpha = 0.05$, $\beta = 0.1$, $\sigma_1/\sigma_2 = 5.83$, $\alpha_M = 0.05$, $e_m = 0.0902$, $\tau_1 = 0.35$, $\tau_2 = 0.06$, $\tau_{ch} = 0.15$.

method	% error	iterations
Non-preconditioned MRII	acc	204
Rank one + 1D fits	2%	158
2-level sym + rank one	acc	174
2-level sym + direct fact	1.5%	160
2-level sym + inverse fact	1.1%	165
1-level sym + direct fact	acc	115
1-level sym + inverse fact	acc	111
Chan preconditioner	acc	145

Table 10: Motion mask, $\alpha = 0.09$, $\beta = 0.01$, $\sigma_1/\sigma_2 = 2$, $\alpha_M = 0.088$, $e_m = 0.0749$, $\tau_1 = 0.5$, $\tau_2 = 0.25$, $\tau_{ch} = 0.28$.

- No differences between the two factorization techniques 5.3.2 and 5.3.3 come up.

Of course in the indefinite case the number of iterations is in general higher and the reconstruction efficiency is lower. Figure 3 refers to the nonseparable indefinite case of Table 8. The blurred image is shown on the left. The reconstructed image with 1-level sym + direct fact preconditioner at 32th iter. is shown on the right.

Conclusions. In summary, we have found that the preconditioners based on the fit technique produce efficient reconstructions at low computational cost. The comparison with the modified circulant Chan preconditioner shows that the fit technique allows to obtain comparable results with fewer iterations (we must point out that the asymptotical cost per iteration of Chan preconditioner is higher). For the nonseparable case, the technique based on the 1-level symmetric two-dimensional fit 5.2(b), followed by one of the factorizations described in 5.3.2 and 5.3.3, outperforms the other fit preconditioners. Also in the separable case this technique can be successfully applied with a possible slight increase of the computational cost (see Section 7). Therefore its use can be suggested when no information on the properties of the mask is available.

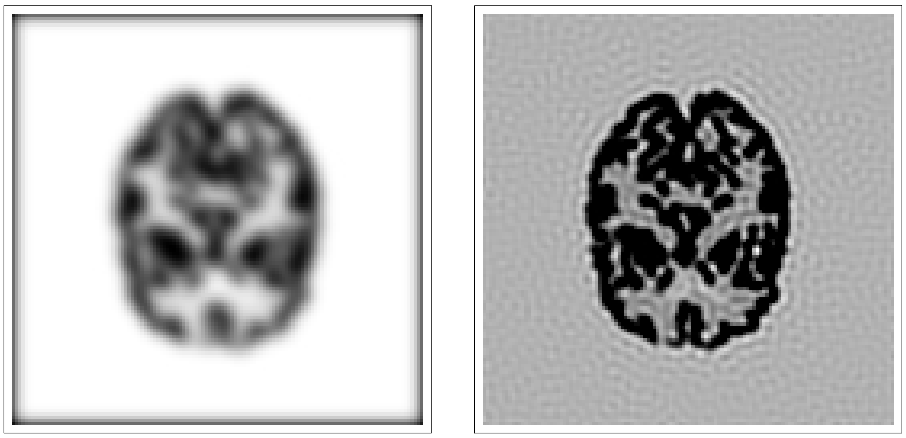

Figure 3: Image blurred with the motion mask of Table 8 on the left, reconstructed image with 1-level sym + direct fact preconditioner at 32th iter. on the right.

References

[1] A. Bacciardi, L. Gemignani, Iterative refinement techniques for the spectral factorization of polynomials, to appear on the Proceedings of the 47th SPIE Annual Meeting, Seattle, WA., 2002.

[2] F.L. Bauer, Beiträge zur Entwicklung numerischer Verfahren für programmesteurte Rechnenanlagen, ii. Direkte Faktorisierung eines Polynoms, *Sitz. Ber. Bayer. Akad. Wiss.*, (1956) 163-203.

[3] M. Bertero, P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, 1998.

[4] D.A. Bini, L. Gemignani, B. Meini, Factorization of analytic functions by means of Koenig's theorem and Toeplitz computations, *Numer. Math.*, **89** (2001) 49-82.

[5] D.A. Bini, L. Gemignani, B. Meini, Computations with infinite Toeplitz matrices and polynomials, *Linear Algebra Appl.*, **343-344** (2002) 26-61.

[6] P. Favati, G. Lotti, O. Menchi, A polynomial fit preconditioner for banded Toeplitz matrices in image reconstruction, *Linear Algebra Appl.*, **346** (2002) 177-197.

[7] J.S. Geronimo, H.J. Woerdeman, Positive extensions and Riesz-Féjer factorization for two-variable trigonometric polynomials (Preprint).

[8] U. Grenander, G. Szegö, *Toeplitz Forms and their Applications*, 2nd edition, Chelsea, N. Y., 1984.

[9] M. Hanke, *Conjugate gradient type methods for ill-posed problems*, (Pitman Research Notes in Mathematics), Longman, Harlow (1995).

[10] M. Hanke, Iterative Regularization Techniques in Image Restoration, in *Mathematical Methods in Inverse Problems for Partial Differential Equations*, Springer-Verlag, 1998.

[11] M. Hanke, J. Nagy, Restoration of atmospherically blurred images by symmetric indefinite conjugate gradient techniques, *Inverse Problems*, **12** (1996) 157-173.

[12] M. Hanke, J. Nagy, R. Plemmons, Preconditioned Iterative Regularization for Ill-Posed Problems, *Numerical Linear Algebra and Scientific Computing*, L. Reichel, A. Ruttan, R. S. Varga eds, de Gruyter, Berlin (1993) 141-163.

[13] E.J. Hoffman, P.D. Cutler, W.M. Digby, J.C. Mazziotta, 3-D phantom to simulate cerebral blood flow and metabolic images for PET, *IEEE Trans. Nucl. Sci.*, **37** (1990) 616-620.

[14] J. Kamm, J. Nagy, Optimal Kronecker product approximation of block Toeplitz matrices, *SIAM J. Matrix Anal. Appl.*, **22** (2000) 155-172.

[15] C. van der Mee, G. Rodriguez, S. Seatzu, LDU Factorization Results for Bi-infinite and Semi-infinite Scalar and Block Toeplitz Matrices, *Calcolo*, **33** (1996) 307-335.

[16] C. van der Mee, G. Rodriguez, S. Seatzu, Spectral factorization of bi-infinite multi-index block Toeplitz matrices, *Linear Algebra Appl.*, **343-344** (2002) 355-380.

[17] M. Miranda, P. Tilli, Asymptotic Spectra of Hermitian Block Toeplitz Matrices and Preconditioning Results, *SIAM J. Matrix Anal. Appl.*, **21**, 3 (2000) 867-881.

[18] W. Rudin, The extension problem of positive definite functions, *Illinois J. Math.*, **7** (1963) 532-539.

[19] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, 1995.

[20] A. H. Sayed, T. Kailath, A Survey of Spectral Factorization Methods, *Numer. Linear Algebra Appl.*, **8** (2001) 467-496.